METHOD FOR THE INDUCTION OF A SPECIFIC RNAi INTO CELLS AND NUCLEIC ACIDS FOR ITS IMPLEMENTATION

[0001] The invention relates to the area of biology and more particularly to the preparation of double-stranded oligonucleotides intended to be used in an RNA interference process (RNAi) in order to induce the degradation of a target RNA.

[0002] Interference RNA, also designated "RNAi", or even co-suppression, has been detected in plants, where it was observed that the introduction of a long, double-stranded RNA corresponding to a gene induces the specific and efficacious repression of the expression of the targeted gene. The mechanism of this interference comprises the degradation of the double-stranded RNA into short duplexes of oligonucleotides of approximately 20 to 22 nucleotides called siRNA's.

[0003] Interference RNA has now been applied to mammals for specifically inhibiting the expression of genes for applications in functional genetics. In fact, siRNA's permit the identification of the function of genes detected by the sequencing of the human genome either in cellular culture models or in animal models and in particular in the mouse. Interference RNA is also useful in the therapeutic area for the treatment or prevention of cancers, infectious diseases and, more generally, of diseases involving a heterologous or homologous mutated gene (S.M. Elbashir, J. Harborth, W. Lendeckel, A. Yalcin, K. Weber and T. Tuschi (2001a), "Duplexes of 21-Nucleotide RNA's Mediate RNA Interference in Cultured Mammalian Cells", Nature 411, 494-498; S. M. Elbashir, J. Martinez, A. Patkaniowska, W. Lendeckel and T. Tuschi (2001b), "Functional Anatomy of siRNA's for Mediating Efficient RNAi in Drosophila melanogaster Embryo Lysate", Embo J. 20 6877-6888).

[0004] siRNA's are short sequences of double-stranded RNA that can be introduced into cells in the form of synthetic oligonucleotides or in the form of vectors permitting the expression of these siRNA's. In this latter approach siRNA is synthesized in the form of a precursor that is a shRNA (short hairpin RNAi), formed by a "rod loop" in which the rod represents the siRNA and the loop any sequence that will be degraded by the cellular RNAase's , which will give birth to the mature siRNA.

[0005] The implementation of siRNA expression vectors has numerous advantages, in particular for applications of functional genetics. It allows double-stranded RNA to be expressed in a stable manner in cells and therefore to more readily inhibit the expression of proteins with a long half-life. In fact, synthetic siRNA's have a half-life of 3 days in mammalian cells. Furthermore, siRNA's are diluted in the course of cellular divisions.

[0006] It also permits long-term effects to be analyzed. On the other hand, it requires the establishment of lines expressing the construction in a stable manner, which presents several disadvantages. In particular, it is necessary to compare stable lines with each other, which is generally difficult to interpret because the cellular lines drift. On the other hand, it is impossible to study the proteins indispensable for the cell since their inhibition will block the proliferation of the cells and will thus prevent the establishing of the stable line. It is therefore indispensable to be able to induce the activity of siRNA. The term "induce the activity of siRNA" denotes being able to block and unblock its activity at will. The term "activity of siRNA" denotes its capacity to recognize and induce the degradation of its target RNA.

[0007] The prior art describes a vector that permits the stable and inducible expression of siRNA. This system is based on the inhibition of the transcription of the precursor of siRNA. In fact, the vector, derived from the pSUPER vector (T. R. Brummelkamp et al,

www.schencexpress.org/March 21, 2002/page 1/10. 1126/science.1068999), contains a known sequence, "the tetracycline operator" ("TET operator"), positioned between the promoter directing the transcription of the siRNA and the sequence coding this siRNA. In the presence of the protein "TET repressor" the latter fixes itself on the "TET operator" and thus blocks the transcripttion of the siRNA. In the presence of doxycycline the latter inhibits the fixation of the "TET repressor" on the "TET operator" and thus permits the transcription of the siRNA. This expression system engenders a relatively elevated background [background noise?] of activity.

[0008] Moreover, the inducible systems based on the promoters functions very poorly in viral vectors because the viral promoters take precedence over the inducible promoters and the background [b. noise?] is very elevated.

[0009] The invention intends specifically to remedy these disadvantages by offering a system that permits the activity of a siRNA to be induced at will.

[0010] This goal is achieved by the present invention by the use of the protein couple TAT-sequence RNA TAR in order to block the activity of a siRNA in mammalian cells.

[0011] The principle of the invention is based on the interaction of TAT/TAR. In fact, the siRNA must be recognized in the cell by a protein complex called RISC complex (RNAi-Induced Silencing Complex) that causes it to undergo a maturation stage and will use it as a guide in order to recognize and degrade the target mRNA. The TAT protein can be used to block this latter stage.

[0012] The TAT protein of VIH plays a crucial part in viral replication. In fact, TAT regulates, as a transcription factor, the replication speed of the virus. TAT is a secreted protein capable of diffusing into non-infected cells and of preparing an environment propitious for viral infection. The TAT protein exists in two forms (71 and 101 and amino acids) coded by two

exons separated by the *env* gene. TAT exercises its function as activator of the promoter of VIH by binding to the RNA TAR sequence present at 5' of all the RNA's of the VIH.

[0013] The linking of TAT and of its transcriptional cellular co-activators to RNA TAR will stimulate the elongation of the transcription by augmenting the processivity of RNA polymerase II.

[0014] The presence of double-stranded RNA's in a cell represents a signal: It involves an RNaseIII called Dicer. This ribonuclease cleaves the long, double-stranded RNA's into small fragments of double-stranded RNA, the siRNA's, with a precise size and structure, in general 21 to 23 nucleotides. The duplex siRNA's then become associated with proteins in order to form a complex called RISC (for "*RNAi-Induced Silencing Complex*"). RISC is therefore a ribonucleo-protein complex containing a molecule of siRNA associated with proteins belonging to the Argonaut family. RISC guides the small siRNA's by virtue of the 5' phosphate end of a simple-stranded siRNA for the specific recognition of the target messenger RNA and catalyses its cleavage; this reaction takes place in the cytoplasm. This terminates in the inhibition of the translation and thus in the inhibition of the expression of the target gene.

[0015] According to the invention the loop of the precursor of the siRNA (shRNA), normally constituted by any sequence, is replaced by a nucleotide sequence coding for at least one nucleo-tide sequence contained in the TAR sequence, minimal and sufficient for being recognized by the TAT protein.

[0016] This sequence referenced in the sequence list submitted in the annex under the num-ber SEQ ID No. 1 is composed of 11 nucleotides and has the following DNA banding pattern [concatenation]: ATCTGAGCTCT (or AUCUGAGCUCU in its RNA form).

[0017] It is also possible to use, in accordance with the invention, a sequence coding for the entire wild or mutated TAR sequence or any fragment of it containing this minimal, non-mutated nucleotide sequence described above. This sequence (DNA or RNA) is named elsewhere in the separate sequence text.

[0018] The separate sequence used in accordance with the invention can be natural or synthetic.

[0019] Thus, in the presence of the TAT protein the latter interacts with the TAR loop on the shRNA, and the interaction of the siRNA protein maturation complex (RISC complex) cannot be realized and the siRNA does not undergo the maturation stages in induced by this complex and therefore remains inactive, that is to say, in the shRNA form.

[0020] In the absence of the TAT protein the TAR loop of shRNA is degraded and the resulting siRNA is assessable to the RISC maturation complex and the degradation process of the target RNA by the RNAi can thus be carried out.

[0021] Therefore, the invention has as subject matter a method for inducing the activity of an RNAi in cells in which:

- The TAT protein and a nucleic acid coding for the sense and antisense sequences of an RNAi of interest are introduced into eukaryote cells, which sense and antisense sequences are separated by a nucleotide sequence comprising at least the sequence referenced under number SEQ I No. 1 in the sequence list supplied as an annex, coding for at least one nucleotide sequence contained in the TAR sequence, minimal and sufficient for being recognized by the TAT protein under conditions in which the nucleic acid is transcribed in RNA, which transcribed separating sequence and the TAT protein form a complex inhibiting the activity of the RNAi of interest;

- The activity of the RNAi is induced by withdrawing the TAT protein.

[0022] The separating sequence can be constituted by any nucleotide sequence provided that it comprises at least the SEQ ID No. 1 sequence. The separating sequence is advantageously constituted by a sequence coding for the complete TAR sequence, wild or mutated, or by a fragment of the latter comprising a sequence coding for the SEQ ID No. 1 sequence.

[0023] According to a particular embodiment of the method of the invention this nucleic acid coding for the sequences comprising the sense and antisense sequences of an RNAi of interest separated by the separating sequence described above can be introduced into the cell in the form of an expression vector, particularly in the form of a plasmid or a viral vector.

[0024] In this embodiment the RNAi of interest is transcribed from the vector coding for shRNA that will give rise to the siRNA after cleavage of the single-stranded separating sequence. Thus, the vector comprises at least one nucleotide sequence coding for the sense and antisense sequences of the siRNA separated by a nucleotide sequence comprising at least the sequence referenced under number SEQ ID No. 1 sequence under the control of a transcription promoter.

[0025] This nucleic acid described above permits the implementation of the method for the induction of the activity of an RNAi in eukaryote cells.

[0026] Moreover, the vector can comprise an antibiotic resistance gene. According to the invention the antibiotic resistance gene is advantageously a neomycin resistance gene.

[0027] The presence of an antibiotic (e.g. neomycin) resistance gene permits the transfected cells to be selected.

[0028] According to the invention, when the cells are cells in culture the TAT protein can be introduced into these cells by a simple culturing of the cells in an environment comprising the TAT protein.

[0029] In fact, an advantage of the present invention resides in the fact that the TAT protein penetrates spontaneously into the cells when they are cultivated in a culture environment containing the TAT protein. The action of siRNA comprising the TAR separating sequence as previously described can therefore be readily blocked by cultivating the cells containing the siRNA in a culture environment containing the TAT protein and to stimulate the activity of the RNAi by cultivating these cells in a culture environment without TAT protein, which results in the degradation of the target RNA by the RNAi.

[0030] Under these conditions the culture environment can contain the TAT protein at a concentration comprised between 0.1 μg/ml and 1.5 μg/ml and very preferentially 1 μg/ml.

[0031] The TAT protein can also be introduced into the cell in the form of an inducible expression vector comprising a nucleotide sequence coding for the TAT protein. In this instance the expression of the protein can be induced, which has as a consequence the inhibiting of the activity of the RNAi and therefore the blocking of the degradation of the target RNA by the RNAi. The degradation of the target RNA by the RNAi can then be induced when the synthesis of the TAT protein is blocked itself.

[0032] According to the invention the cells transfected with the nucleic acid of the invention are mammalian cells. The method applies equally well to the transfection of cells in culture as well as directly in the animal.

[0033] The invention permits genes especially human genes, to be analyzed in a reliable manner from a functional point of view in cells in culture or in animals, especially in mice.

[0034] The invention also relates to cell or a cell line in which a nucleic acid in accordance with the invention as previously described was introduced.

[0035] The invention also relates to animals comprising cells in which the nucleic acid in accordance with the invention as previously described was introduced.

[0036] Finally the invention has compositions, especially pharmaceutical compositions, as subject matter comprising as active substance at least one nucleic acid in accordance with the invention as previously described or cells containing this nucleic acid such as previously described, possibly associated in the composition with a compatible excipient.

[0037] Other advantages and features of the invention will appear from the following examples in which reference is made to the attached drawing in which figure 1 represents the inhibition of the GFP marker by RNAi.

Example 1: Construction of the plasmid pTATOF-siRNAGFP coding for an RNAi whose target is the messenger RNA coding for Green Fluorescent Protein (GFP):

[0038] This plasmid is constructed from the pSUPER plasmid, permitting the constitutive expression of siRNA and described by Brummelkamp et al.

[0039] The DNA sequence (synthetic DNA oligonucleotide synthesized in a hand-tailored manner) containing in order the sequences SEQ ID No. 2 (coding for the sense siRNA) - SEQ ID No. 1 (coding for the separating sequence) - SEQ ID No. 3 (coding for the antisense siRNA) is introduced immediately upstream from the H1 promoter of the pSuper plasmids at sites Bgl II in 5', and Hind III in 3'. This thus yields the expression vector pTATOF-siRNAGFP.

(SEQ ID NO. 2) coding for the sense siRNA: 5' GCAAGCTGACCCTGAAGTTC 3' 3' CGTTCGACTGGGACTTCAAG 5'
Sequence coding for the separating sequence (SEQ ID No. 1) 5' ATCTGAGCTCT 3' 3' TAGACTCGAGA 5'
(SEQ ID NO. 3) coding for the antisense siRNA: 5' GAACTTCAGGGTCAGCTTGC 3' 3' CTTGAAGTCCCAGTCGAACG 5'

[0040] Thus, the sequences coding for the sense and antisense siRNA are separated by a minimal TAR loop coded by the separating sequence called SEQ ID No. 1 coding for the minimal sequence recognized by the TAT protein (SEQ ID No. 4).

Coding for the siRNA:
5' 3'
GCAAGCTGACCCTGAAGTTCATCTGAGCTCTGAACTTCAGGGTCAGCTTGC
CGTTCGACTGGGACTTCAAGTAGACTCGAGACTTGAAGTCCCAGTCGAACG

Example 2: Inhibition of the expression of GFP by the siRNAGFP expressed by the vector of example 1:

[0041] Mammalian cells COS-7 are transfected with polyfect (Qiagen) with 4 μg of expression vectors of the siRNA (pTATOF-siRNAGFP, part A, or pSuper siRNAGFP without TAR loop, part B) as well as an expression vector of Green Fluorescent Protein or GFP (500 ng). Sixty hours after the transfection a Western blot was performed from the total extracts using an antibody directed against the GFP (Santa cruz) or cellular tubulin (Sigma) in order to evaluate the quantity of proteins used for this test.

[0042] The cells are cultivated in a DMEM (Gibco) environment containing 10% of fetal bovine serum in the presence (wt) or absence (-) of a vector for the expression (0.5 μg) of TAT protein or of its mutants, co-transfected (V. Bres et al., *Nat. Cell Biol.* Aug. 2003; 5(8): 754-61).

[0043] A control is performed with cells that did not receive the expression vector of siRNA (pTATOF-siRNAGFP or pSuper siRNAGFP) cultivated in the absence of TAT protein. Likewise, controls are performed with a vector for the expression of the mutated TAT protein (incapable of bonding the TAR sequence) in the presence of the vector for the expression of siRNA (pTATOF-siRNAGFP or pSuper siRNAGFP).

[0044] Figure 1 shows the results of this experiment:

Part A:

In the absence of TAT protein and of pTATOF-siRNAGFP, the GFP is present in the cells (column 1).

In the absence of TAT protein and in the presence of the of pTATOF-siRNAGFP vector, the GFP is degraded (column 2).

In the presence of wild TAT protein and of the of pTATOF-siRNAGFP vector the GFP is present in the cells (column 3) in a quantity comparable to that present in the cells that received neither TAT protein nor pTATOF-siRNAGFP vector (column 1) and is therefore not degraded.

In the presence of mutated TAT protein and of the of pTATOF-siRNAGFP vector the GFP is degraded in the cells (columns 4, 5).

Part B:

In the absence of TAT protein and of pSupersiRNAGFP vector (without TAR sequence) the GFP is present in the cells (column 1).

In the absence of TAT protein and in the presence of the pSupersiRNAGFP vector the GFP is degraded (column 2).

In the presence of wild or mutated TAT protein and of the pSupersiRNAGFP the GFP is degraded in the cells (columns 1, 4 and 5).

[0045] This demonstrates that the TAT protein recognizes the separating sequence intercalated between the sense and antisense sequences of siRNA and blocks its activity when a mutated TAT protein incapable of recognizing the separating sequence has no effect on the siRNA that appears perfectly functional.